

Proteine Bolognese ¹

R. Cubarsi^a, E. Vázquez^b, A. Villaverde^b

^aDept. de Matemàtica Aplicada IV, Univ. Politècnica de Catalunya, Barcelona

^bInstitut de Biotecologia i Biomedicina, Dept. de Genètica i Microbiologia, Univ. Autònoma de Barcelona, Bellaterra

1. Introducción

- Para reforzar la asimilación de conceptos matemáticos y estimular el aprendizaje, sería conveniente introducir en los planes de estudios de carreras científico-técnicas una asignatura de Modelización Matemática, con ejemplos ligados a otras disciplinas (biología, economía, física etc.), donde se manejaran conceptos específicos de otras asignaturas de matemáticas, así como conocimientos transversales de más asignaturas.
- En el ejemplo que seguidamente se estudia se puede dar una solución analítica, a partir de métodos de Ecuaciones Diferenciales. También se puede usar software matemático para las representaciones gráficas y para los métodos de estimación de parámetros y errores. Si el modelo inicial, a posteriori, se quisiera ajustar más a los datos, sería también ineludible el uso de otras técnicas numéricas y, quizás, de simulación.
- Los métodos numéricos para la estimación de los parámetros del modelo –estimación por mínimos cuadrados, en nuestro caso– y los métodos de estimación y propagación de errores son omnipresentes en todas las áreas de la ciencia. Los primeros se estudian en asignaturas de Cálculo Numérico, mientras que los segundos son propios de asignaturas de Probabilidades y Estadística. Los alumnos utilizan estas técnicas desde los primeros cursos, en las prácticas de laboratorio de física, química, biología, etc.
- Por lo tanto, la Modelización Matemática es una buena oportunidad para profundizar en todos estos conceptos y métodos.

2. Experimento

- Se estudia cómo entra en una célula de mamífero, que posee un núcleo, una proteína verde fluorescente llamada GFP, que al llevar fusionado un péptido altamente catiónico, forma pequeñas partículas de 20 nm de diámetro. Dicha proteína se ha ingeniarizado [1] para penetrar células de mamífero y transportar genes u otros fármacos a su núcleo, con finalidades terapéuticas.
- Las células de mamífero son más o menos esféricas, y cuando están en cultivo son un poco más planas, de forma que, para simplificar, supondremos que adoptan forma cilíndrica, con el núcleo centrado en el eje de simetría. Cuando se sumerge una célula en un medio que contiene esta proteína, se detecta la presencia de fluorescencia, primero en el citoplasma solamente, y después en el citoplasma y en el núcleo.
- La detección se hace a partir de secciones ópticas de la célula, de grosor constante y suficientemente pequeño, con un microscopio láser llamado confocal, pudiendo escoger el radio de la sección a fin de incluir sólo el núcleo, o bien citoplasma y núcleo conjuntamente. De las imágenes digitalizadas se obtienen distintas mediciones para cada sección y, sumando, para todo el volumen considerado. Por ejemplo, número de píxeles, luminosidad de cada píxel, área o volumen del total de píxeles, valores máximo, mínimo y medio de la luminosidad, densidad lumínica por unidad de superficie o de volumen, etc. Con los datos obtenidos para diferentes tiempos, se desea obtener un modelo dinámico para la penetración de la GFP en el citoplasma y en el núcleo.
- La variable escogida para estudiar la evolución temporal de la GFP ha sido la fluorescencia absoluta $\phi(t)$, medida en unidades de fluorescencia, bien sea de toda la célula, del núcleo, o por diferencia, del citoplasma. En la Tabla de la Fig. 2 se muestran los datos para una de las células: fluorescencia absoluta (en unidades de fluorescencia) de toda la célula ϕ_T , del núcleo ϕ_N y del citoplasma ϕ_C , para cada tiempo t (min.).

¹Jornadas i-Math de Modelización Matemática, Universitat Politècnica de València, València, 2010.

3. Modelo

• Para la célula completa, la fluorescencia total $\phi(t)$ (se omite en este caso el subíndice T) en función del tiempo es casi nula al principio del experimento, y luego es una cantidad siempre positiva y continuamente creciente, que tiende a estabilizarse. Por similitud a la dinámica de poblaciones, parece lógico tomar el modelo de la ley logística, dado por la ecuación diferencial:

$$\frac{1}{\phi} \frac{d\phi}{dt} = p\phi + q; \quad q > 0, \quad p < 0 \quad (1)$$

con p y q constantes. Definiendo $F(t) = \ln \phi(t)$, la tasa de variación o aumento relativo de la fluorescencia viene dada por $F'(t) = \frac{d \ln \phi}{dt}$.

• Al inicio, cuando $\phi \approx 0$, F' vale q , que es la máxima tasa de fluorescencia. Después, la tasa F' baja progresivamente hasta anularse, cuando la fluorescencia se estabiliza. La máxima fluorescencia es $L = -\frac{q}{p}$. Estos son los dos únicos puntos de equilibrio del sistema: $\phi = 0$, inestable, y $\phi = L$, estable. A diferencia del modelo poblacional, no tiene sentido un valor $\phi > L$. La solución de la Eq. 1 es inmediata y su discusión se encuentra en muchos textos básicos de ecuaciones diferenciales [2].

• Para determinar los parámetros del modelo a partir de un conjunto de m valores $(t, \phi(t)); t = t_1, \dots, t_m$, se usará la expresión discretizada

$$F'(t_i) = p\phi(t_i) + q \quad (2)$$

siendo necesario hacer la estimación numérica de la derivada temporal $F'(t)$. Aunque en este caso las medidas temporales se han tomado a intervalos regulares, en otros casos no es así. Por lo tanto, para conseguir una mejor aproximación, hacemos una estimación de la derivada de la función $F(t)$ en tres puntos. Así, en un tiempo intermedio $t_i = t_2, \dots, t_{m-1}$ obtenemos [3]

$$F'(t_i) = \frac{(t_{i-1} - t_i)^2 F(t_{i+1}) - ((t_{i-1} - t_i)^2 - (t_{i+1} - t_i)^2) F(t_i) - (t_{i+1} - t_i)^2 F(t_{i-1})}{(t_{i+1} - t_i)(t_{i-1} - t_i)(t_{i-1} - t_{i+1})} \quad (3)$$

• El siguiente paso es resolver el sistema lineal sobredeterminado por el método de mínimos cuadrados, en los $n = m - 2$ puntos donde se tiene calculada $F'(t)$. Escribimos dicho sistema en la forma matricial $\mathbf{y} = \mathbf{G}\mathbf{x}$, donde $\mathbf{x} = (p, q)^T$. Puesto que los datos iniciales son mediciones únicas y no llevan asociada una medida de error, asignaremos a cada una de ellas un error experimental σ_y , tal que el error cuadrático medio del ajuste de regresión sea $\sigma_y^2 = \frac{1}{n} |\mathbf{y} - \mathbf{G}\mathbf{x}|^2$. Finalmente, la estimación de los errores de los parámetros se calcula, como es usual [4], a partir de la matriz de covarianzas de los errores experimentales, $\mathbf{V}_x = (\mathbf{G}^T \mathbf{G})^{-1} \sigma_y^2$. En el Cuadro 1 se muestran los estimadores óptimos de los parámetros p, q para la célula completa y el núcleo, así como sus errores (desviación estándar).

• El modelo propuesto sería básicamente aplicable a toda la célula y al núcleo, por encontrarse en situaciones parecidas: la proteína entra y la fluorescencia aumenta hasta llegar a la saturación. En cambio, el citoplasma podría no mostrar necesariamente una luminosidad creciente, pues su fluorescencia podría descender a medida que aumenta la del núcleo. Sin embargo, observamos que en muchos casos sí que es efectivamente posible usar el mismo modelo para aproximar los parámetros del citoplasma. Por tanto, también se listan los parámetros del citoplasma en el Cuadro 1.

• En la Fig. 1 se muestran las rectas de regresión: tasa de fluorescencia F' en función de la fluorescencia ϕ , según la Eq. 1, para toda la célula (izquierda), citoplasma (centro) y núcleo (derecha). Los cuadrados corresponden a datos experimentales y la línea continua al ajuste de regresión.

• El modelo no explica la zona entre líneas verticales discontinuas.

Componente	p	q	L
Total	$-1,51 \cdot 10^{-8} \pm 1,23 \cdot 10^{-9}$	$0,82 \cdot 10^{-2} \pm 0,44 \cdot 10^{-3}$	$5,43 \cdot 10^5 \pm 5,31 \cdot 10^4$
Citoplasma	$-2,11 \cdot 10^{-8} \pm 2,89 \cdot 10^{-9}$	$0,63 \cdot 10^{-2} \pm 0,53 \cdot 10^{-3}$	$2,97 \cdot 10^5 \pm 4,80 \cdot 10^4$
Núcleo	$-4,05 \cdot 10^{-8} \pm 2,69 \cdot 10^{-9}$	$1,02 \cdot 10^{-2} \pm 0,46 \cdot 10^{-3}$	$2,53 \cdot 10^5 \pm 2,03 \cdot 10^4$

Cuadro 1: Parámetros p, q y L con su error (desviación estándar).

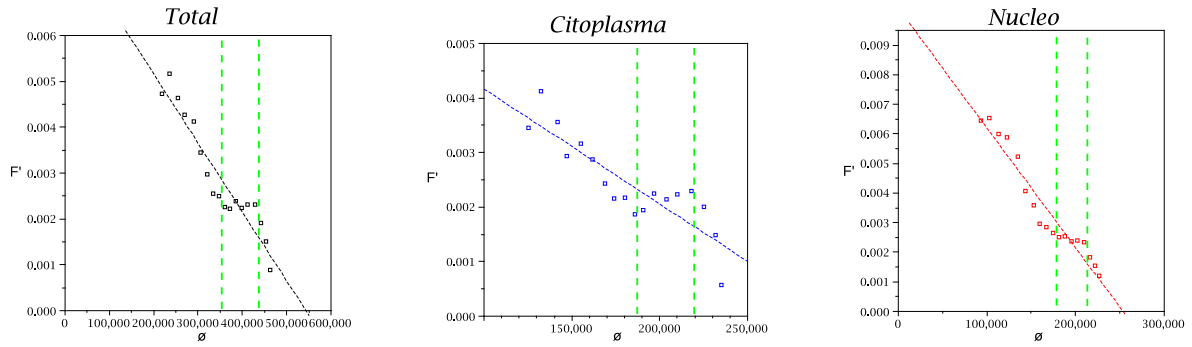


Figura 1: Tasa de fluorescencia F' en función de la fluorescencia ϕ .

4. Discusión

- Aunque el modelo proporciona una primera aproximación al proceso estudiado, con una disminución progresiva de la tasa de fluorescencia a ritmo casi constante a medida que la luminosidad total va aumentando hasta saturarse, observamos que los gráficos de la tasa de fluorescencia tienen una tendencia característica intercalada en la línea de regresión: en el recorrido descendente de la tasa de fluorescencia, hay un período de estabilidad del ritmo de fluorescencia, incluso con un ligero aumento de dicha tasa (Fig.1, entre las líneas verdes discontinuas). Ello es debido a un proceso que no se realiza de forma totalmente diferenciable.
- En la Fig. 2 se muestra la evolución temporal de la fluorescencia. Las curvas, para toda la célula (en negro) y para el núcleo (en rojo), son de tipo sigmoide, como corresponde a la solución de la Eq. 1. Para el citoplasma (en azul), obtenido como diferencia entre las anteriores, también se asemeja a una función sigmoide, aunque podría no ser así si los límites de fluorescencia entre el total y el núcleo fueran similares.
- A partir de las anteriores figuras podemos verificar si esta “anomalía” del modelo se produce, o no, simultáneamente en el citoplasma y en el núcleo. En la Fig. 1 marcamos las fluorescencias inicial y final de esta zona con la línea discontinua verde. Luego, en las curvas temporales de la Fig. 2, leemos los tiempos correspondientes a estas fluorescencias. Los límites de la región seleccionada son comunes para el núcleo, el citoplasma y la célula completa. La simultaneidad es muy evidente en el extremo de la derecha de la zona señalada, para todas las partes de la célula (Fig. 1). El inicio de la anomalía, en cambio, es el mismo para la célula completa y para el citoplasma, con un cambio brusco de comportamiento de la tasa de fluorescencia, pero se produce de forma mucho más suave en el núcleo.
- En consecuencia, parece que la membrana nuclear retiene GFP en el citoplasma y, posteriormente, su ruptura reproduce una sobrefluorescencia similar en el núcleo. No obstante, mientras esto sucede en el núcleo, todavía hay un aumento de la tasa de fluorescencia en el citoplasma, en lugar de un descenso brusco, como sería lógico. Esto es posible explicarlo por una entrada súbita de GFP en el citoplasma, mayor que el flujo que sale hacia el núcleo. Es decir, por una ruptura o penetración súbita de la membrana del citoplasma.
- La confirmación de esta hipótesis la proporciona el gráfico del comportamiento global de la célula (Fig. 1, Total), donde también se produce un sensible aumento de la tasa de fluorescencia durante este intervalo de tiempo.
- Concluimos pues, que no sólo es útil la correcta modelización de un experimento, sino que, de un modelo que no se ajusta totalmente a los datos, también se pueden deducir resultados interesantes, precisamente porque nos proporciona información relevante sobre la desviación del patrón que se ha adoptado.
- Apuntamos las siguientes posibles causas:
 - (a) Irrupción brusca de GFP en el citoplasma y en el núcleo, por debilitación o ruptura de sus membranas, lo cual produciría una tendencia a la estabilización, o aumento de la tasa de fluorescencia, en la parte posterior a la membrana.
 - (b) Obstrucción del paso de GFP por la membrana nuclear, lo cual produciría un aumento de la tasa de fluorescencia del citoplasma.

t [min]	ϕ_T	ϕ_N	ϕ_C
90	203109	84084	119025
105	217723	92602	125121
120	234113	102097	132016
135	254372	112749	141623
150	269190	122292	146898
165	289247	134582	154665
180	304716	143141	161575
195	320808	152154	168654
210	333268	159449	173819
225	346376	166432	179944
240	359224	173695	185529
255	370714	180322	190392
270	384155	187412	196743
285	398401	194708	203693
300	411066	201264	209802
315	427259	209381	217878
330	440822	215982	224840
345	452704	221297	231407
360	461427	226326	235101
375	465128	229619	235509

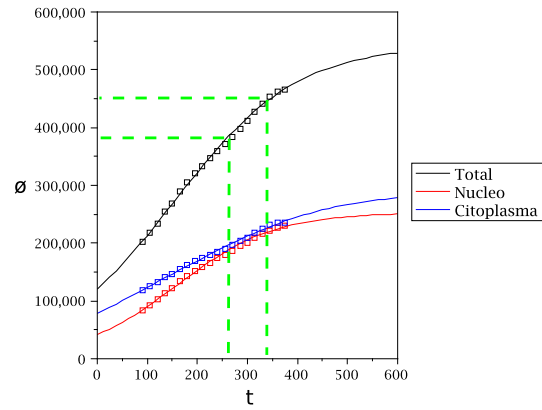


Figura 2: Fluorescencia en función del tiempo.

5. Conclusión

- Los nuevos planes de estudios, de acuerdo con las directrices del Proceso de Bolonia [5], iniciado en 1999 con la finalidad de unificar del Espacio Europeo de Educación Superior (EEES), han consolidado la estructura cuatrimestral de la mayoría de las asignaturas, forzando una revisión a fondo de sus contenidos. Se contempla un incremento de horas lectivas no presenciales –trabajo personal del alumno– para compensar la reducción de horas lectivas presenciales que supone la nueva reestructuración. En consecuencia, las asignaturas de matemáticas, impartidas en grados de titulaciones técnicas y experimentales, pueden únicamente contener los conocimientos estrictamente necesarios para poder realizar las otras asignaturas de la carrera.
- Para que la nueva metodología dé sus frutos, hay que plantear un temario muy optimizado, en el que cada tema entronque con temas de otras asignaturas. Así, se pueden aprender unos contenidos específicos de cada asignatura, y practicar o reforzar los contenidos específicos de otras asignaturas.
- El tiempo nunca sobra y los temarios, heredados de planes de estudios anteriores, tienden a ser demasiado extensos. Especialmente los seres humanos aprendemos a base de imitar y repetir un mismo procedimiento, hasta conseguir hacerlo propio. Aquí podemos recordar aquello de “Ein Mal ist kein Mal” (literalmente, una vez es ninguna vez). Para asimilar los contenidos, es necesario repetir del derecho y del revés los mismos procesos de aprendizaje. Y si no ¿cómo se aprenden las tablas de multiplicar, los modos y tiempos de verbos, o se adquiere destreza con un instrumento musical?
- La Modelización Matemática es un buen pretexto para aplicar y revisar la mayor parte de los conocimientos adquiridos en otras asignaturas, en particular de matemáticas y, además, una forma eficaz de estimular el aprendizaje, dado que muestra la utilidad de lo que se estudia.

Referencias

- [1] E. Vázquez *et al.*, Nanomedicine (Lond), 5(2):259-68, 2010.
- [2] M.W. Hirsch, S. Smale, *Differential Equations, Dynamical Systems, and Linear Algebra*. Academic Press, New York, 1974.
- [3] D. Kincaid, W. Cheney, *Numerical Analysis*. Brooks/Cole Publishing, Belmont, CA, 1991.
- [4] R. Barlow, *Statistics, a guide to the use of statistical methods in the physical sciences*. John Wiley & Sons, Chichester, 1989.
- [5] http://europa.eu/legislation_summaries/education_training_youth/lifelong_learning/c11088_es.htm